informationresearch
VOL. 16 NO. 3, SEPTEMBER, 2011

Measuring the dimensions of serendipity in digital environments

Lori McCay-Peet
Dalhousie University, 6100 University Avenue, Halifax, Nova Scotia, Canada B3H 3J5
Elaine Toms
Information School, University of Sheffield, Regent Court, 211 Portobello, Sheffield, United Kingdom, S1 4DP

Abstract

Introduction. Serendipitous information retrieval is the perhaps inevitable consequence of immersion in an information-rich environment. Just how well chance encounters are supported, however, within these environments varies and one of the challenges to the development of tools and systems to facilitate serendipity is measuring how well they achieve this goal. This research developed a scale to measure dimensions of serendipity identified in prior research.
Method. Participants (N=123) browsed an experimental information search system for twenty minutes with no a priori task and responded to a twenty-item survey questionnaire. Items were derived from the serendipity dimensions of a physical library setting devised by Björneborn.
Analysis. Exploratory factor analysis using the Principal Component Analysis as the method of extraction was carried out on the data. The analysis was undertaken using the SPSS statistical package.
Results. Five factors were extracted representing core elements of support for serendipity in a digital environment: *enabled connections*, *introduced the unexpected*, *presented variety*, *triggered divergence*, and *induced curiosity*. In addition, four of the original twenty items were eliminated from the survey.
Conclusions. While the physical dimensions of serendipity do map on to the digital dimensions, it is unknown whether there are additional dimensions of serendipity not present in the physical environment.

Introduction

Unintended but auspicious discoveries have been credited with ground-breaking global advancements in medicine and science. At the same time, less tangible breakthroughs have aided advancements and innovation within organizations and on a much smaller scale, growth and development of individuals in personal and work lives. Yet our knowledge of serendipity primarily comes from our observation of it in the physical, tactile world of science, rather than in conceptual and often non-observable world of social sciences where the discovery may be the unplanned use of a new method that has unintended consequences, the uncovering of a long forgotten cultural gem, or, the chance connection between two disparate ideas that leads to a new theory. These potential outcomes of serendipitous discovery come from interaction with content that is textual, audio or graphic embedded within a system or from exchange with other people. In this research we focus on the former; serendipity in interactive information systems. Notably, current systems may not provide the rich environment necessary to enable serendipity and instead tend to provide planned paths of navigation from one discrete information object to another with the intent of removing the apparent irrelevant content or shielding the user from what the system perceives is not useful to the user.

When does a system enable, support and facilitate serendipity? A number of innovations have been introduced as we discuss in the next section. But rather than taking a technology-centred view of the problem, we instead start with known dimensions of serendipity identified by previous research to create a tool to measure a system's propensity to facilitate serendipity. The objective of this research was two-fold: to assess the applicability of these dimensions to information systems and, secondly, to examine selected characteristics or indicators of serendipity and assess them in an ecologically valid user study.

Previous work

There has been a swell of research that explores the concept of serendipity in various domains such as everyday life (Rubin *et al.* 2010; Williamson 1998), education (Nutefall and Ryder 2010; Sawaizumi *et al.* 2010), and interdisciplinary scholarship (Foster and Ford 2003). The characteristics associated with the serendipity-prone person have also been examined (Erdelez 1999; Heinström 2006) to determine why it is that some people appear more likely to experience serendipity than others. People come to digital information environments with diverse individual differences, motivations, and knowledge which impact their interactions with the environment and their overall experience. Serendipity, however, is not only influenced by individual differences and information activities of the individual, but also the characteristics of the environment in which the individual is immersed (Björneborn 2008; Erdelez and Rioux 2000).

Interdisciplinary scholars in Foster and Ford's (2003) study indicated that their own attitudes about serendipitous information encounters and their strategies for dealing with them had a perceived effect on how well these encounters could be exploited but they also suspected that the characteristics of the environment such as the classification schemes of libraries played a significant role.

Studies have generally tested tools or methods of introducing unexpected and interesting content. Efforts have been made to support serendipity in digital environments, for example, through the development of a context and activity-aware recommender system embedded in mobile leisure guides (Bellotti *et al.* 2008), information retrieval software (MAX) that collects potentially serendipity-inducing information and send it to users by e-mail (Campos and Figueiredo 2002), an ambient intelligence agent (Mitsikeru) within an internet browser (Beale 2007) and a suggested articles tool in an information search system (Toms 1999; Toms and McCay-Peet 2009). Techniques for deriving serendipity-inducing content such as personalization (André *et al.* 2009) have also begun to be tested. Measurement of the effectiveness of these techniques and systems has generally relied on judgments by participants of the novelty, relevance, and interestingness of the results or suggestions provided (André, *et al.* 2009; Bellotti, *et al.* 2008; Campos and Figueiredo 2002) as well as how the tool was actually used (i.e., support for the task at hand or divergence from the task) (Toms and McCay-Peet 2009).

In addition, while the aforementioned studies have selected aspects of serendipity to test and explore in a digital environment, little is known about which precipitating conditions are actually related to serendipity in a digital environment. Precipitating conditions (Cunha 2005) or principles of serendipity (Fine and Deegan 1996) are those conditions that, when present, increase the chances of serendipity. While they have been hypothesized to include temporal happenstance, social networking and active learning, in a digital environment there are likely to be many potential characteristics that may increase our opportunity for serendipity.

One study examined the impact of design features of the physical library space on convergent and divergent information behaviour (Björneborn 2008). Björneborn conceptualized serendipity as an event precipitated by divergent information behaviour and the design of the *library interface* (human, physical and digital). Over a period of several months, the information activities that patrons used to find resources in public libraries were observed. A portion (n=118) of those patrons was invited to engage in a short interview to collect self-reported data on their information behaviour. Follow-up think-aloud sessions with eleven of the 118 patrons were conducted in which the users walked through the library with a researcher and reflected upon their information behaviour and what triggered their attention in the physical library space.

Ten dimensions in the physical library that facilitate serendipity were identified and these primarily revolve around the design elements of the library space (Björneborn 2008):

1. *Unimpeded and direct access* to library resources to ensure that serendipity is not restricted by lack of access to resources.
2. *Rich diversity* of resources, activities, and physical spaces that will help spark an individual's interests.
3. *Curiosity-invoking display* of library resources.
4. *Striking contrasts* between spaces and resources within the library spaces that will draw an individual's attention to the rich diversity.
5. *Pointers* such as signs to remind individuals of potential topics of interest.
6. *Imperfections* within the library interfaces such as mis-shelved books that may lead an individual to stumble upon an unexpected and interesting resource.
7. *Cross contacts,* which is the juxtaposition of divergent topics, activities, resources, or physical spaces.
8. *Multi-reachability,* which is the ability of individuals to reach library spaces through a variety of routes or pathways.

9. *Explorability,* which is the encouragement of the individual to move freely around the library, to explore and follow their curiosity.
10. *Stopability,* which is the ability for individuals to stop and explore materials at will through, for example, the provision of a space to sit or a place to lay down what they were carrying to examine another resource.

While interactions with computers in the library were observed, details of these interactions are at a high level; more research is needed to determine whether the same serendipity dimensions exist in the digital environment (Björneborn 2008). Several of the serendipity dimensions combined (e.g., *multi-reachability*, *stopability* and *explorability*) appear at face value to have a counterpart in the digital realm in the form of orientation and wayfinding tools that support serendipity and exploratory search. However, the characteristics of these dimensions and the behaviour they encourage may cause distraction and disorientation in a digital environment. Unlike the *gestalt nature* of physical information spaces such as libraries, there is a limit to the amount of information that can be revealed on a screen (Rice 1988) and when they are in the middle of a task, individuals have a tendency to want to stay on task (Erdelez 2004; Toms and McCay-Peet 2009). While users need pathways to divergent content and time to absorb that content, they also need to get back on track after digressions from their original path (Toms 2000). *Imperfections* is also a difficult dimension to translate in a digital environment where imperfections are viewed as frustrating usability issues rather than opportunities for serendipity. Similarly *unimpeded and direct access* seems to be first and foremost related to usability; it is important to all digital information interactions, not just those designed to support serendipity. *Rich diversity*, *curiosity invoking display*, *striking contrasts*, *pointers* and *cross contacts* are similar in their endeavour to catch an individual's attention and spark their interest through some manipulation of the physical library environment. In a digital environment this takes the form, for example, of personalized results (André, *et al.* 2009), a suggestion box with semi-related content recommendations (Toms 1999) and visually augmented hyperlinks to interesting and unexpected content (Beale 2007).

The digital environment's proper manipulation and design is the goal of much of the current serendipity research discussed here. The belief is that if we can understand enough about the process of serendipity and the conditions that precipitate it, we can create an environment that will encourage and support serendipity (Björneborn 2010). Two questions emerge from this endeavour: 1) what characteristics of digital information spaces or environments (i.e., design of information spaces) facilitate chance encounters; and 2) how can we measure how well digital environments facilitate chance encounters through system design? This paper focuses on the latter question by applying dimensions of serendipity derived from the physical environment (Björneborn 2008) to a digital information space and endeavours to answer the following questions:

1. Do the dimensions of serendipity in the physical environment map to dimedimensionsserendipity in the digital environment?
2. How can we assess the degree to which a digital envienvironmentilitates serendipity?

Methods

To assess the dimensions of serendipity, we developed an initial scale adapted from Björneborn's (2008) ten dimensions of the physical library that may facilitate serendipity and tested the scale on a set of users of a version of Wikipedia with a custom-designed interface. This was part of a larger study that is examining non-goal-based use of an information system in tandem with measures of curiosity.

Design of the study

The study invited a volunteer community to examine a version of Wikipedia viewing anything they wished for a period of twenty minutes. Before being assigned the task, several measures related to curiosity were assessed and, after the task, they were assigned several questions regarding their found information and the serendipity scale. Responses to the serendipity items were on a five-point Likert scale of 'strongly disagree' to 'strongly agree'. No experimental variables were formally tested.

Task

Participants were given a scenario, a role play, in which they were asked to imagine that they are in an office with about twenty minutes to spend. They were given a laptop and asked to browse whatever they wish, including retaining items for future use or making notes.

System

Participants used the wikiSearch system (described in Toms, *et al.* 2009) which included the content of Wikipedia which was used for the INEX (INitiative for the Evaluation of XML Retrieval) 2006 Interactive Track and a custom-designed interface which contained a single display to represent the search process. The interface included three physical columns:

1. the left column supported the task and contained in addition to the task assignment, a Notes space for memorable information, and a BookBag to retain items of interest;
2. the middle pane supported the search task and included a box for query entry, a set of results with abbreviated titles that additionally provided mouse-over access to the complete snippet, and a history of all page views and submitted queries; and
3. the extreme right side contained the page display. It also included a Suggested Items list for other items that may be related somewhat to the page on display, created by resending the first paragraph in the page displayed as query and returning the top rank titles (not including the one on display).

The system was intended to retain as many visual cues as possible about what the user had done or examined, with significant flexibility about what could be re-examined.

Instrument

While multiple instruments were used in this study, we focus on that used in the analysis presented here: the serendipity scale.

Through an iterative process, seven research assistants and doctoral students were given Björneborn's (2008) ten dimensions, each with corresponding statements extracted by the researchers from the intent of the dimension. They rated whether each statement captured the essence of the dimension and were invited to provide additional statements if they perceived that the presented statements were incorrect or incomplete. This process started with 115 statements across all ten dimensions and the list was amplified by a further seventeen statements. Over two more iterations, the list stabilised in a process that resulted in the twenty items in Table 1.

As previously suggested, some of Björneborn's (2008) dimensions were difficult to adapt to a digital environment. In operationalizing a conceptually written dimension, it became clear that considerable overlap existed among items in the set, making it difficult assign a particular statement to a particular dimension. This was the case, for example, with *striking contrasts* and *pointers*, both of which prompted items related to *attention-grabbing*. On the other hand, it was difficult to develop any items for *imperfection*. *Imperfection* may work in a physical environment, but conceptualizing it in a digital environment led to items that appeared to describe usability problems. In the end, we elected to include statements that captured the essence of the concept, serendipity, while not attending to the need to provide multiple items. We recognized that in fact each dimension may not resolve to a factor in the final analysis.

Serendipity dimensions (Björneborn 2008)	Scale items
1. Unimpeded and direct access	S1 I was able to explore anything that interested me when using the system
2. Rich diversity	S2 I was able to see information in a range of formats S3 I was able to examine a variety of topics S4 I explored many topics that normally I do not examine
3. Curiosity-invoking display	S5 I wanted to click on things to see where they would take me
4. Striking contrasts	S6 Unexpected words and phrases caught my eye S7 Unexpected visual features of the system caught my eye
5. Pointers	S8 Unexpected words and phrases sparked my thinking
6. Imperfections	n/a
7. Cross contacts	S9 I found something interesting on pages that had unexpected content S10 The system enabled me to make connections between different topics S11 The system presented content in ways that invited me to explore across topics S12 Exploring one topic unexpectedly led me to other S13 I stumbled upon unexpected topics

8. Multi-reachability	S14 I could find topics in several alternative ways
9. Explorability	S15 The system encouraged me to browse and explore
10. Stopability	S16 I found myself pausing to look at things more closely S17 The system encouraged me to stop and explore S18 I could return to topics that I had explored earlier S19 I could easily explore many topics without getting lost S20 I could follow-up on interesting associations between topics or ideas

Table 1: Serendipity dimensions applied as items relating to a digital environment

Participants

One hundred and twenty-four participants volunteered for the study. Of these, 123 were identified as 'valid participants'. One participant's data set was identified as an extreme multivariate outlier and removed from the analysis. The participants (N = 123) were primarily university students (N = 106, 86.2%) between the ages of 18 and 26 (N = 98, 79.7%). There was a fairly even representation of males (N = 67, 54.5%) and females (N = 56, 45.5%). Participants were frequent users of Wikipedia; 99 participants (80.4%) indicated that they used Wikipedia more than once or twice a week. Of those, 34 (27.6%) participants indicated that they used Wikipedia one or more times a day. The majority of participants (N=119, 96.7%) also indicated using web search engines one or more times daily.

Procedure

Notices for recruitment were posted around the University and advertised on electronic mailing lists. Participants self-selected and no formal random sampling was done. Participants *dropped* into a room that was set up for the study. The room contained ten tables and laptops scattered about in an informal arrangement.

Participants were introduced to the study by one of the five research assistants, read the consent form and had their questions answered, and then were logged into the system that guided them through the process.

On completion, they were paid an honorarium of $20 for the nearly one hour of time taken to do the study.

Data analysis

Data were loaded into SPSS 17.0. All univariate and multivariate assumptions were examined. Linearity, homoscedasticity and multimulti-collinearity all found to be satisfactory and no missing values were detected. One multivariate outlier was found and removed and all results are reported based on the remaining 123 participants. First, we examined descriptive statistics and inter-correlations for the twenty variables as well as testing for sampling adequacy. Next exploratory factor analysis was conducted in multiple iterations to test for factors present within the set of items and finally Cronbach's Alpha was used to assess each sub-scale emerged from the exploratory factor analysis.

Results

Table 2 contains the descriptive statistics and inter-correlations of all twenty variables. As illustrated, all correlated with one of more other items at the .05 to .01 levels, and were retained for the subsequent analyses. Exploratory factor analysis was conducted on the 20 items using the Principal Component Analysis as the method of extraction with orthogonal rotation (Varimax) with Kaiser normalisation.

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
S1	3.8	1.2	1																			
S2	3.3	1.0	.30**	1																		
S3	4.2	0.6	.45**	.25**	1																	
S4	3.3	1.2	.11	.19*	.10	.1																
S5	4.0	0.9	.17	.24**	.07	.18*	1															
S6	3.4	1.1	-.05	.11	.02	.09	.31*	1														
S7	2.7	1.1	.17	.44**	.21*	.22*	.13	.17	1													
S8	3.6	1.0	.09	.12	.14	.08	.09	.48**	.15	1												
S9	3.8	0.9	.06	.10	.19*	.24*	.14	.24**	.30**	.20*	1											
S10	3.9	0.9	.49**	.09	.33**	.10	.18	.01	.14	.30**	.40	1										
S11	3.8	0.8	.38**	.16	.39**	.02	.21	.23**	.23*	.27**	.10	.53**	1									
S12	4.2	0.8	.26**	.24*	.45*	.21*	.12*	.25*	.28*	.19*	.34**	.19	.36**	1								
S13	3.7	1.0	-.04	.12	.18*	.45**	.14	.23*	.08	.18*	.43**	.07	.14	.38**	1							
S14	3.8	1.0	.26*	.23**	.32**	.17	.19	.08	.16	.20*	.20*	.39**	.43**	.38**	.18	1						
S15	3.9	0.8	.31*	.49**	.37**	.12	.24*	.24**	.23*	.19*	.18*	.29**	.47**	.47**	.20*	.31**	1					
S16	3.9	0.9	.06	.10	.13	.15	.16	.06	.13	.12	.16	.19*	.16	.19*	.05	.11	.13	.1				
S17	3.8	0.9	.42*	.28**	.30**	.23**	.28**	.20*	.17	.14	.27**	.34**	.36**	.44**	.31**	.26**	.49**	.42**	1			
S18	4.5	0.7	.16	-.06	.31**	.04	.09	.05	-.20*	.22*	.24**	.26**	.18*	.05	.19*	.24**	.10	.24*	.25**	1		
S19	4.1	0.9	.10	.06	.15	-.07	.20*	.05	.09	.13	.22*	.34**	.35**	.17	.02	.28*	.23**	.12	.26**	.32**	1	
S20	4.0	0.9	.32**	.04	.36**	.07	.23*	.14	.13	.36**	.13*	.66**	.48**	.27*	.01	.31**	.23*	.39**	.40**	.33**	.31**	1
Note. N = 123.,*p <0.05,**p <0.01.																						

Table 2: Summary of inter-correlations, means, and standard deviations of items

The study sample met the minimum recommended number of participants (i.e., five to ten) for each variable (Kass and Tinsley, 1979, in Field 2009). The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.76, indicating that the present set of data is suitable for Principal Component Analysis; in general, values between 0.70 and 0.80 are considered good (Kaiser 1974, as cited in Field 2009). All Kaiser-Meyer-Olkin values for individual items were >0.61 which is above the acceptable limit of 0.5 (Field 2098). Similarly, Bartlett's test of sphericity was significant ($X^2 = 790.25$, df=190, p<0.001), indicating sufficient correlations between variables to proceed with the analysis.

Four iterations of exploratory factor analysis were performed. Because the sample size was less than 200, the cut-off value for interpretation of 0.45 (20% of the variance) was selected (Meyers *et al.* 2006). Item loadings were interpreted as excellent with loadings of 0.7, very good with loadings of 0.63, good with loadings of 0.55, fair with loadings of 0.45, and poor with loadings of 0.32 (Comrey and Lee 1992, as cited in Meyers *et al.* 2006).

The first iteration of Exploratory Factor Analysis resulted in six factors. Item loadings ranged from 0.49 (S5) to 0.83 (S13). Factor 6 contained only one item (S19). One variable (S20) was complex, loading strongly on both factors one and four. Therefore, for the next iteration these two items were removed. Five factors were extracted in the second iteration. Item loadings ranged from 0.53 (S15) to 0.83 (S13). Item S12 loaded strongly on factors one and two, and was removed in the next iteration. In the third iteration, items loaded on five factors and loadings ranged from 0.49 (S15) to 0.84 (S6). None of the items above the 0.45 cut-off loaded on more than one factor. Cronbach's alpha, calculated for each of the five sub-scales, ranged from 0.26 (Factor 3) to 0.77 (Factor 1). S18, which had a high negative loading (-0.59), was found to cause a substantial decrease in α and was therefore removed (Field 2009).

In the fourth iteration, five factors were extracted among the remaining sixteen items; no complex items were found. Therefore, these items were retained in the final solution (See Table 3). The five factors account for 60.8% of the total variance which, according to Tabachnick and Fidell (in Meyers *et al.* 2006) suggests a '*good solution*' where a robust solution accounts for at least 50% of the variance. Communalities (h^2) were relatively high, ranging from .44 to .77. Cronbach's alpha for each of the five sub-scales, ranged from 0.55 (Factor 5) to 0.75 (Factor 1). Of the five factors, only Factor 1 had an acceptable alpha level (>0.70). This may be due in part to the small number of items contained in Factors 2-5 (i.e., two to three items for each factor). All five factors, however, were retained because of their conceptual importance, and low reliabilities were taken into account in the study findings. Items are ordered and grouped by size of loading to facilitate interpretation. Factor loadings less than 0.45 (20% of variance) are represented as zeros. Interpretive labels are suggested for each of the five factors in the footnote and discussed below.

Item #	Item	Factors*					
		F_1	F_2	F_3	F_4	F_5	h^2
S10	The system enabled me to make connections between different topics	**0.78**	0.00	0.00	0.00	0.00	0.66
S11	The system presented content in ways that invited me to explore across topics	**0.72**	0.00	0.00	0.00	0.00	0.66
S3	I was able to examine a variety of topics	**0.68**	0.00	0.00	0.00	0.00	0.56
S1	I was able to explore anything that interested me when using the system	**.68**	0.00	0.00	0.00	0.00	0.56
S14	I could find topics in several alternative ways	**0.60**	0.00	0.00	0.00	0.00	0.44
S13	I stumbled upon unexpected topics	0.00	**0.84**	0.00	0.00	0.00	0.73
S9	I found something interesting on pages that had unexpected content	0.00	**0.67**	0.00	0.00	0.00	0.52
S4	I explored many topics that normally I do not examine	0.00	**0.67**	0.00	0.00	0.00	0.53
S2	I was able to see information in a range of formats	0.00	0.00	**0.85**	0.00	0.00	0.77
S7	Unexpected visual features of the system caught my eye	0.00	0.00	**0.69**	0.00	0.00	0.53
S15	The system encouraged me to browse and explore	0.00	0.00	**0.51**	0.00	0.00	0.55
S6	Unexpected words and phrases caught my eye	0.00	0.00	0.00	**0.84**	0.00	0.77
S8	Unexpected words and phrases sparked my thinking	0.00	0.00	0.00	**0.77**	0.00	0.68
S16	I found myself pausing to look at things more closely	0.00	0.00	0.00	0.00	**0.76**	0.61
S17	The system encouraged me to stop and explore	0.00	0.00	0.00	0.00	**0.64**	0.68
S5	I wanted to click on things to see where they would take me	0.00	0.00	0.00	0.00	**0.58**	0.49
Eigenvalues		4.28	1.83	1.33	1.12	1.10	
Percentage of variance		26.77	11.40	8.34	7.42	6.85	
Cronbach's a		0.75	0.63	0.65	0.64	0.55	
*Factor labels: F_1 Enabled connections; F_2 Introduced the unexpected; F_3 Presented variety; F_4 Triggered divergence; F_5 Induced curiosity							

Table 3: Summary of Factor Loadings and Communalities (h^2) for the four iterations of Principal Component Analysis (N = 123)

Factor 1: Enabled connections

This factor accounted for 26.7% of the variance and consists of five items. Item loadings on this factor ranged from 0.60 to 0.78. These items related to the manner in which the system invited connections between topics (S10), exploration or browsing across topics (S1, S3, S11, 14). Since the items related to making connections between topics and exploration, this factor was labelled *enabled connections*.

Factor 2: Introduced the unexpected

This factor accounted for 11.4% of the variance and consisted of three items. Item loadings ranged from 0.67 to 0.84. These items related to encountering unexpected topics or content (S9, S13) and topics not normally explored (S4) therefore this factor was labelled *introduced the unexpected*.

Factor 3: Presented variety

This factor accounted for 8.3% of the variance and consisted of three items. Item loadings ranged from 0.51 to 0.85. These items related to the variety of system content and display (S2 and S7) as well as the facilitation of varied or diverse behaviours such as exploration and browsing (S15). Thus, *presented variety* was determined to be an appropriate label for factor three.

Factor 4: Triggered divergence

This factor accounted for 7.42% of the variance and consisted of two items. Item loadings in this factor were 0.77 and 0.84. Both items (S6, S8) indicated that the system in some way sparked or triggered their attention and initiated divergent thinking and behaviour. This factor, therefore, was labelled *triggered divergence*.

Factor 5: Induced curiosity

The final factor accounted for 6.85% of the variance and consisted of three items. Item loadings ranged from 0.58 to 0.76. The items in this factor related to the inducement of deeper exploration or consideration of information encountered (S16, S17) and curiosity-teasing triggers (S5). The label selected for this factor was *induced curiosity*.

Discussion

Not all applications have the potential to support serendipity, but there is an increasing need to enhance our ability to be innovative with creative or creative-facilitating information systems. A step toward that goal is the development of systems that enable serendipity to occur. But we need to know when a system is meeting that goal. Using a set of characteristics developed by Björneborn (2008) of the serendipity-enabling dimensions of a physical library space, this study adapted these dimensions to a digital environment. In the course of the analysis Börneborn's (2008) original ten dimensions of serendipity were reduced to five factors: *enabled connections, encountered unexpected, presented variety, triggered divergence*, and *induced curiosity*.

Factor 1, labelled *enabled connections*, reflects aspects of several characteristics of Björneborn's (2008) physical serendipity dimensions, namely, unimpeded and direct access, rich diversity, cross contacts, all of which facilitate connections or associations between disparate information and resources. Making connections or bisociations is a key component of the process of serendipity (Cunha 2005). Koestler (1964) originated the term *bisociation* to describe a creative association made between two disparate pieces of information or skills. An example of a bisociation is Gutenberg observing the presses in action during a wine harvest in Rhineland and subsequently designing a printing press capable of mass production (Simonton 1995). In an information system, this could mean finding an unexpected piece of information that takes the conceptualization of your problem or task in a new direction (Foster and Ford 2003).

Encountered the unexpected, factor 2, is related to Björneborn's (2008) original *rich diversity* and *cross contacts* dimensions. Diversity and cross contacts are perhaps bound to lead one to the unexpected and unexpected is perhaps the umbrella or over-arching description of the experience of serendipity. Something about the experience must be unexpected or unanticipated. The example of the journalist who stumbles upon the juxtaposition of two previously unrelated concepts in a Wikipedia entry is one example of encountering the unexpected in a digital environment (McCay-Peet and Toms 2011). All of the systems that have attempted to support serendipity discussed in this paper have attempted in some way to introduce unexpectedness through, for example, the provision of content that is intended to be novel and unexpected (Campos and Figueiredo 2002).

Factor 3, *presented variety*, whose items were derived from Björneborn's (2008) dimensions of *rich diversity, striking contrasts* and *explorability* points to the importance of the provision of variety in content and the importance of systems drawing attention to this variety and allowing ways to explore it. Creativity research indicates that '*Bombardment by the diversity of surrounding events eventually primes that train of thought that can connect the disparate parts of the desired Koestlerian bisociation*' (Simonton 1995: 483). This suggests that environments that create opportunities for viewing divergent materials and enabling interesting juxtapositions may not only support serendipity, but potentially prime for it.

The fourth factor, *triggered divergence*, can be traced to Björneborn's (2008) dimensions of *striking contrasts* and *pointers*. It is an aspect of serendipity that has been grappled with in systems such as Mitsikeru (Beale 2007), an ambient intelligence system that attempts to call attention to interesting or surprising content (Web pages) through visual cues. While its efficacy in enabling serendipity was not established, the importance of a system's ability to trigger divergence is evident due to the limits of human attention. Fine and Deegan (1996: 436) write that for ethnographers, serendipity is linked to their '*ability to make sense of seemingly chance events; that is, these scholars were able to 'keep their wits about them,' finding, in the rush of on-going events, meanings and opportunities that might escape others*'. That '*rush of on-going events*' aptly describes a digital search experience in which the world is literally at your fingertips. Because of the limitations of working memory, the attentional portion of short-term memory, we can only allocate a certain amount of our cognitive resources to a scene; we attend to a limited amount of information (Wood, *et al.* 2006). Because information encounters contain new pieces of information, our capacity to attend to them and make the necessary associations and bisociations for serendipity, is greatly affected by our working memory. Interestingly, '*super-encounterers*' appear reluctant to expose themselves to environments such as the Web where they feel unable to cope with the associated information overload (Erdelez 2000), perhaps the multiple divergences it triggers.

The final factor that emerged in this study, *induced curiosity*, relates to Björneborn's (2008) dimensions of *curiosity-invoking display* and *stopability* in which curiosity is both ignited and stoked. This element of serendipity is reflected in the development of a physical, serendipity card system (Sawaizumi, *et al.* 2007) to help students capture stray ideas and thoughts and think more deeply about them. Students were asked to record interesting observations they encountered, indicating a hypothesis and theme as well as the who, what, when, where, and how. The *induced curiosity* factor underlines perhaps more so than the other four the importance of the individual in the success or failure of a serendipity-enabling environment. The individual must be curious about what is being displayed and become actively engaged; the individual is not a passive observer. While Björneborn focused on the extrinsic or physical rather than intrinsic characteristics of the individual (e.g., personality, interests, motivations) of divergent behaviour within the library, he noted that internal factors have role in serendipity. All of the dimensions in some way have a potential impact on the way patrons feel, behave, think, and are motivated as they make their way through the library, encouraging them to stop and explore.

The paper set out to resolve two questions: A. Do the dimensions of the serendipity in the physical environment map to dimensions of serendipity in the digital environment? And B. How can we assess the degree to which a digital environment facilitates serendipity? In the case of the first question, we have identified which of the dimensions of serendipity in a physical environment map to those in a digital environment. In answer to the second question, we assessed how well a digital environment facilitated serendipity by using the provision of a scale. We used a scale because serendipity cannot be objectively assessed by a third party. Only the individual who is facing the potential of a serendipitous moment, a veritable auspicious chance encounter, can assess whether dimensions of serendipity are present. By deploying a more comprehensive scale we can capture the experience of serendipity in a more holistic manner than asking questions that simply relate to the nature of the information encounter as unexpected, interesting, or useful.

This study took place in the context of a larger study that is looking more broadly at the exploratory process used in examining information without a specific goal in mind. In the past, we have discovered that participants in laboratory studies are less likely to step outside the bounds of the study (see Toms and McCay-Peet 2009), and enable the potential for having a truly serendipitous experience. They tend to follow instructions. With no goal, the research design gave the participants the right to choose the direction and pathway through the information space, and thus enable the potential for a serendipitous experience. We know that serendipity is an interaction of some sort between what the user knows (or wants to know), or has experienced and the environment in which the user is immersed (McCay-Peet and Toms 2011). This design provided the environment for that to occur.

Conclusion and future work

This study identified five core dimensions of serendipity that are present in digital environments. To do so, we operationalized a set of dimensions of a physical library developed from a naturalistic study by Björneborn (2008). In the process of operationalizing Björneborn's ten dimensions, it became clear that these dimensions are not discrete and, in the end, did not map neatly on to the dimensions of serendipity in a digital environment. However, the overlapping nature of the ten dimensions suggests that it is plausible that these dimensions are not discreet in a physical world either.

Scale construction is an iterative process. Further development of this scale is required before it can be used to assess the degree to which a digital environment facilitates serendipity. The items used in this scale were developed solely through the interpretation of serendipity dimensions in a *physical*/library setting. Future studies will attempt to expand on these to determine whether there are dimensions unique to a digital environment. Are five factors the complete set? For example, in a physical library setting, the serendipity experienced was that of finding a resource or topic of interest; in a digital environment it may be more about information interactions at the content level. Examination of information or knowledge-based examples of serendipity may indicate other potentially useful scalar items that reflect serendipity within a digital environment. By expanding the items initially tested together with a wider participant base, we can come closer to the creation of a scale that will measure whether a system is meeting its goal of providing an environment that encourages serendipity. Future studies will augment and refine the instrument developed in this study.

Acknowledgements

Research was supported by grants to Toms from the National Centres of Excellence GRAND project, the Social Sciences and Humanities Council of Canada, the Canadian Foundation for Innovation and the Canada Research Chairs Program, while she was at Dalhousie University, Halifax, Nova Scotia, Canada.

About the authors

Lori McCay-Peet is a Ph.D. student in the faculties of Management and Computer Science at Dalhousie University. Her primary research interests include understanding serendipity and measuring its support in order to develop tools to support serendipity within information-rich systems. Her research is in part funded through a Joseph-Armand Bombardier CGS Doctoral Scholarship, a Canadian national scholarship granted by the Social Sciences and Humanities Research Council, and the GRAND National Centres of Excellence. She can be contacted at: mccay@dal.ca
Elaine Toms is Professor in Information Science, Information School, University of Sheffield, and former Canada Research Chair in Management Informatics, Dalhousie University, Halifax, Nova Scotia, Canada. Her research focuses at the intersection of human computer interaction and information retrieval with a particular interest in user-centric evaluation of information systems. She has a long-time interest in the immersive experiences in information that enable chance encounters. She can be contacted at: e.toms@sheffield.uk

References

- André, P., Teevan, J. & Dumais, S.T. (2009). From x-rays to silly putty via Uranus: serendipity and its role in web search. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems*. (pp. 2033-2036). New York, NY: ACM Press. Retrieved 8 August, 2011 from http://research.microsoft.com/pubs/79633/chi09-serendip.pdf (Archived by WebCite® at http://www.webcitation.org/60mbYQpio)
- Beale, R. (2007). Supporting serendipity: using ambient intelligence to augment user exploration for data mining and web browsing. *International Journal of Human-Computer Studies*, **65** (5), 421-433.
- Bellotti, V., Begole, B., Chi, E.H., Ducheneaut, N., Fang, J., Isaacs, E., *et al.* (2008). Activity-based serendipitous recommendations with the Magitti mobile leisure guide. In *Proceedings of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems*.
- Björneborn, L. (2008). Serendipity dimensions and users' information behaviour in the physical library interface. *Information Research*, **13**(4), paper 370. Retrieved 19 June 2011 from http://informationr.net/ir/13-4/paper370.html (Archived by WebCite® at http://www.webcitation.org/60maSB29Y)
- Björneborn, L. (2010). Design dimensions enabling divergent behaviour across physical, digital, and social library interfaces. In Thomas Ploug, Per Hasle & Harri Oinas-Kukkonen, (Eds.). *Persuasive technology: 5th International Conference, PERSUASIVE 2010, Copenhagen, Denmark, June 2010, Proceedings* (pp. 143-149). Berlin: Springer. (Lecture Notes in Computer Science 6137)
- Campos, J. A. & Figueiredo, A. D. (2002). Programming for serendipity. In *Proceedings of the 2002 AAAI Fall Symposium*. (pp. 48-60). Menlo Park, CA: AAAI Press. Retrieved 8 August, 2011 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.12.1856&rep=rep1&type=pdf (Archived by WebCite® at http://www.webcitation.org/60mhdbYo2)
- Cunha, M.P.E. (2005).Serendipity: why some organizations are luckier than others. Lisbon: Universidade Nova de Lisboa, Faculdade de Economia. (FEUNL Working Paper No. 472.) Retrieved 8 August, 2011 from http://fesrvsd.fe.unl.pt/WPFEUNL/WP2005/wp472.pdf (Archived by WebCite® at http://www.webcitation.org/60mhquf6E)
- Erdelez, S. (1999). Information encountering: it's more than just bumping into information. *Bulletin of the American Society for Information Science*, **25**(3), 25-29.
- Erdelez, S. (2000). Towards understanding information encountering on the Web. *Proceedings of the ASIS Annual Meeting*, **25**(37), 363-371.
- Field, A.P. (2009). *Discovering statistics using SPSS*, London: Sage Publications.
- Fine, G.A. & Deegan, J.G. (1996). Three principles of serendip: insight, chance, and discovery in qualitative research. *International Journal of Qualitative Studies in Education*, **59**(3), 321-340.
- Heinström, J. (2006). Psychological factors behind incidental information acquisition. *Library & Information Science Research*, **28**(4), 218-228.
- Koestler, A. (1964). *The act of creation*. New York, NY: Macmillan.
- McCay-Peet, L. & Toms, E.G. (2011). *Exploring the precipitating conditions of serendipity*, Poster presented at the 2nd Annual Graphics, Animation and New Media (GRAND) NCE Conference 2011, Vancouver, BC, May 12-14, 2011.
- Meyers, L.S., Gamst, G. & Guarino, A.J. (2006). *Applied multivariate research: design and interpretation*.. Thousand Oaks, CA: Sage Publications.
- Nutefall, J.E. & Ryder, P.M. (2010). The serendipitous research process. *The Journal of Academic Librarianship*, **36**(3), 228-234.

- Rice, J. (1988). Serendipity and holism. The beauty of OPACs. *Library Journal* **113**(3), 138-141.
- Rubin, V.L., Burkell, J. & Quan-Haase, A. (2010). Everyday serendipity as described in social media. *Proceedings of the 73rd ASIS&T Annual Meeting*, **47**, Article 152. Retrieved 8 August, 2011 from http://www.asis.org/asist2010/proceedings/proceedings/ASIST_AM10/submissions/409_Final_Submission.pdf (Archived by WebCite® at http://www.webcitation.org/60miZrX48)
- Sawaizumi, S., Katai, O., Kawakami, H. & Shiose, T. (2007). Using the concept of serendipity in education. In *KICSS 2007 : The Second International Conference on Knowledge, Information and Creativity Support Systems: Proceedings of the Conference, November 5-7, 2007*. Retrieved 8 August, 2011 from: https://dspace.jaist.ac.jp/dspace/bitstream/10119/4087/1/13.pdf (Archived by WebCite® at http://www.webcitation.org/60mikN4hu)
- Simonton, D.K. (1995). Foresight in insight? A Darwinian answer. In R.J. Sternberg & J.E. Davidson (Eds.), *The nature of insight* (pp. 465-494). Cambridge, MA: MIT Press.
- Toms, E.G. (1999). What motivates the browser? In Thomas D. Wilson and David K. Allen, (Eds.) *Exploring the Contexts of Information Behaviour: Proceedings of the 2nd International Conference on Research in Information Needs, Seeking and Use in Different Contexts*. (465-494). London: Taylor Graham.
- Toms, E.G. (2000). Understanding and facilitating the browsing of electronic text. *International Journal of Human-Computer Studies*, **52**(3), 423-452.
- Toms, E.G. & McCay-Peet, L. (2009). Chance encounters in the digital library. In M. Agosti, J. Borbinha, S. Kapidakis, C. Papatheodorou & G. Tsakonas (Eds.), *ECDL'09 Proceedings of the 13th European conference on Research and advanced technology for digital libraries,* (pp. 192-202). Berlin: Springer-Verlag. (Lecture Notes in Computer Science 5714).
- Toms, E.G., McCay-Peet, L. & Mackenzie, R.T. (2009a). wikiSearch: from access to use. In M. Agosti, J. Borbinha, S. Kapidakis, C. Papatheodorou & G. Tsakonas (Eds.), *ECDL'09 Proceedings of the 13th European conference on Research and advanced technology for digital libraries,* (pp. 27-38). Berlin: Springer-Verlag. (Lecture Notes in Computer Science 5714).
- Williamson, K. (1998). Discovered by chance: the role of incidental information acquisition in an ecological model of information use. *Library & Information Science Research* **20**(1), 23-40.

How to cite this paper

McCay-Peet, L. & Toms, E. (2011). "Uses and gratifications: measuring the dimensions of serendipity in digital environments" *Information Research*, **16**(3) paper 483. [Available at http://InformationR.net/ir/16-3/paper483.html]

Find other papers on this subject